Parasitology Research (2019) 118:2753–2766
https://doi.org/10.1007/s00436-019-06420-9

IMMUNOLOGY AND HOST-PARASITE INTERACTIONS - REVIEW

Plasmodium helical interspersed subtelomeric family—an enigmatic piece of the *Plasmodium* biology puzzle

Vikash Kumar[1] · Ankita Behl[2] · Rachana Sharma[1] · Aanchal Sharma[1] · Rachna Hora[1]

Received: 14 June 2019 / Accepted: 2 August 2019 / Published online: 15 August 2019

Abstract

Plasmodium falciparum (*Pf*) refurbishes the infected erythrocytes by exporting a myriad of parasite proteins to the host cell. A novel exported protein family '*Plasmodium* Helical Interspersed Subtelomeric' (PHIST) has gained attention for its significant roles in parasite biology. Here, we have collected and analysed available information on PHIST members to enhance understanding of their functions, varied localization and structure-function correlation. Functional diversity of PHIST proteins is highlighted by their involvement in *Pf*EMP1 (*Pf* erythrocyte membrane protein 1) expression, trafficking and switching. This family also contributes to cytoadherence, gametocytogenesis, host cell modification and generation of extracellular vesicles. While the PHIST domain forms the hallmark of this family, existence and functions of additional domains (LyMP, TIGR01639) and the MEC motif underscores its diversity further. Since specific PHIST proteins seem to form pairs with *Pf*EMP1 members, we have used in silico tools to predict such potential partners in *Pf*. This information and our analysis of structural data on a PHIST member provide important insights into their functioning. This review overall enables readers to view the PHIST family comprehensively, while highlighting key knowledge gaps in the field.

Keywords Malaria · *Plasmodium falciparum* · PHIST · *Pf*EMP1 · Cytoadherence · MEC motif

Introduction

Malaria is a major human health problem in several developing countries, which is caused by *Plasmodium* species (*P. falciparum*, *P. vivax*, *P. ovale*, *P. malariae* and *P. knowlesi*) (World Health Organization 2016). Amongst these, *Plasmodium falciparum* (*Pf*) causes the most lethal form of the disease due to occurrence of severe complicated malaria. *Plasmodia* reside in mature red blood cells (RBCs) that are devoid of a nucleus, cell organelles and the major histocompatibility complex (MHC) conferring the parasite an ability to survive and evade the host immune response. Owing to the absence of erythrocytic trafficking machinery, parasites establish their own apparatus inside the host cell for nutrient uptake and protein export (Kirk and Saliba 2007). Several *Pf* proteins, collectively called the 'exportome', are translocated from within the parasite confines to the infected red blood cell (iRBC) membrane and cytosol to facilitate host cell remodelling (Maier et al. 2009), possibly increasing disease severity (Haldar and Mohandas 2007). Modifications include appearance of knob-like structures on the iRBC membrane, formation of ion channels for the uptake of extracellular nutrients and enhanced host cell membrane rigidity (Crabb et al. 1997; Soni et al. 2016). These alterations in parasitized host cells are accomplished by several exported protein families including *Pf*EMP1, PHIST and *Pf*2TM (Maier et al. 2008). Several exported proteins are translocated to their destinations via an intermediate membranous structure in parasitized RBCs called Maurer's clefts (MC) (Lanzer et al. 2006). Also, many exported proteins (~400) carry a *Plasmodium*

Vikash Kumar and Ankita Behl contributed equally to this work.

Section Editor: Dana Mordue

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00436-019-06420-9) contains supplementary material, which is available to authorized users.

✉ Rachna Hora
rachnahora@gmail.com

[1] Department of Molecular Biology and Biochemistry, Guru Nanak Dev University, Amritsar, Punjab, India

[2] Department of Biotechnology, Guru Nanak Dev University, Amritsar, Punjab, India

export element (PEXEL motif) to assist their transport across the parasitophorous vacuole membrane via the translocon (PTEX) (Hiller et al. 2004; Marti et al. 2004).

PHIST (*Plasmodium* helical interspersed sub-telomeric) is one of the novel exported protein family that is believed to have diverse roles in trafficking of *Pf*EMP1 proteins, their stabilization in the iRBC membrane and modulating iRBC membrane rigidity and cytoadherence (Maier et al. 2008; Prajapati and Singh 2013; Sargeant et al. 2006). The knob localized *Pf*EMP1 protein family has been long known as an important effector of cytoadherence, the phenomenon by which iRBCs adhere to host endothelial cells in the microvaculature (Su et al. 1995). This obstructs blood flow to vital organs such as the brain, liver and placenta resulting in malaria-related complications such as cerebral and placental malaria (MacPherson et al. 1985; Pongponratn et al. 1991; Sharma and Shukla 2017).

Sequence characteristics and domain organization of PHIST proteins

The PHIST family comprises ~14% of all PEXEL-positive proteins (Sargeant et al. 2006; Warncke et al. 2016). Phist genes have a two-exon structure, where the first exon encodes a signal sequence and the second codes for a PEXEL motif and a 'Phist' domain (Sargeant et al. 2006; Warncke et al. 2016). Sargeant et al. reported 72 paralogs of this family in *P. falciparum*, 39 in *P. vivax* and 27 in *P. knowlesi* amongst the human infecting *Plasmodia (*Sargeant et al. 2006*)*. The ~160 amino acid long PHIST domain is unique to *Plasmodium* and comprises of four conserved alpha helices. Members of the PHIST family are further classified into the sub-families PHISTa, PHISTb and PHISTc based on the presence and location of several conserved tryptophan residues (Sargeant et al. 2006). While PHISTa harbours only two tryptophan residues within the PHIST domain, the others have four. The PHIST sub-familes differ in their domain architecture and sequence complexity where PHISTa proteins are very short consisting of only a signal sequence, an export motif and the PHIST domain. PHISTb proteins show more variability in their C-terminal portion and include seven RESA (ring-infected surface antigen)-like proteins where the PHIST domain is fused with a DnaJ domain. PHISTc proteins form the most diverse group, members of which have been reported to bind ATS of *Pf*EMP1 and *Pf* skeleton binding protein 1 (*Pf*SBP1) (Kumar et al. 2018). Interestingly, the primary sequence of PHIST domains is not very well conserved. One-to-one BLAST search using the sequence of the PHIST domain of PFI1780w as the query with other members of the PHIST family shows less than 30% identity. However, PHIST family proteins have 99–100% sequence identity amongst orthologs in different *Plasmodium* species.

Apart from the presence of a conserved PHIST domain in all family members, 55 PHIST proteins harbour an additional ~60 amino acid long TIGR01639a domain. This domain is *Plasmodium* specific and is always located within the PHIST domain at its N-terminus. However, the functional role of this domain is still obscure.

Members of PHIST family are essential for parasite survival

Several members of the PHIST family have been reported to be essential for parasite survival, which highlights their significance in malaria biology. We used PhenoPlasm database to identify the total number of essential PHIST proteins (Table S1) (Sanderson and Rayner 2017). A total of 4 and 22 PHIST members were found to be essential by Maier et al. and Zhang et al., respectively (Maier et al. 2008; Zhang et al. 2018). All four members reported by Maier et al. were shown to be non-essential by Zhang et al., and a few PHIST proteins (PF3D7_0731300, PF3D7_110250, PF3D7_1102500) identified as essential by Zhang et al. were earlier reported to be non-essential by Maier et al. This discrepancy in the results obtained in the two studies is likely to be a result of different methodologies used. While Maier et al. used a large-scale gene knockout strategy along with functional screens to identify essential exported proteins involved in host cell remodelling, Zhang et al. utilized *piggybac* transposon insertion sites to attain saturation-level mutagenesis to identify essential genes (Maier et al. 2008; Zhang et al. 2018).

PHIST members show differential localization

Different PHIST family members have been reported to localize to distinct subcellular regions. Two PHISTc proteins (PFI1780w and PF11_0503) and nine PHISTb proteins (PFE1605w, PFD0080c, PFD0080c, PFD1170c, PFI1770w, PF11_0037, PFL2535w, PF14_0730 and PFA0110w) were exported to the iRBC surface (Kumar et al. 2018; Oberli et al. 2014; Tarr et al. 2014). PHISTb protein PFE1605w was seen to co-migrate with PfEMP1 within iRBCs before being localized at the knobs (Oberli et al. 2014). However, PFI1780w was present uniformly along the iRBC membrane and was completely absent from the knobs (Oberli et al. 2014). PHISTc proteins PFD1140w and MAL7P1.172 (*Pf*PTP2–*Pf*EMP1 trafficking protein)] localized to Maurer's clefts (Kumar et al. 2018; Maier et al. 2008). While MCs seem to be the final destination of PFD1140w [18], PfPTP2 was also found in iRBC generated extracellular vesicles (EV) (Regev-Rudzki et al. 2013) that carry several other PHIST proteins (PHISTa-PFD0090c, PHISTb-PFD0080c, PFD1170c and PHISTc-PF08_0137, PFI1780w) (Abdi et al.

2017). iRBC generated EVs have been reported to perform important roles like communication amongst parasitized erythrocytes, immune modulation and parasite development (Regev-Rudzki et al. 2013). Other subcellular locations for PHIST proteins are J-dots (*Pf*PHIST_0801-PHISTc) (Zhang et al. 2017) and parasitophorous vacuolar membrane (PVM) (PF3D7_0402000-PHISTa) (Parish et al. 2013). Presence of PHIST members in different subcellular and extracellular localizations suggests existence of divergent roles for this protein family that have been entailed in the following sections.

PHIST members play diverse roles in malaria biology

PHIST members bind *Pf*EMP1 to perform varied functions

The phenomenon of cytoadherence is mostly effected by the *Pf*EMP1 family of proteins via interaction of their large extracellular portions with a broad range of host endothelial receptors including CD36, intercellular adhesion molecule 1 (ICAM 1), chondroitin 4-sulfate (CSA) and P-selectin (Sherman et al. 2003). The extracellular regions of *Pf*EMP1 consist of multiple variable duffy-binding-like (DBL) domains, cysteine-rich interdomain regions (CIDR) and C2 domains, while their cytoplasmic acidic terminal segments (ATS) are conserved (Sherman et al. 2003; Smith et al. 2000a, b). A PHISTb protein PFB0080c was reported to bind in vitro with the extracellular domain of *Pf*EMP1 (VAR2CSA-PFL0030c) and regulate expression of knob-associated Hsp40 (Goel et al. 2014). Here, PFB0080c was proposed to regulate *var* gene expression and cytoadherence properties of iRBCs based on enhanced surface expression of VAR2CSA in PFB0080c knockouts. Also, PFB0080c knockouts maintained VAR2CSA expression for several generations suggesting a role for this protein in controlling switching of *var* genes (Goel et al. 2014).

Some members of the PHIST family viz. PFI1780w, PFE1605w, PFD1140w and PF11_0503 showed in vitro interaction with the ATS domain of *Pf*EMP1 (PF08_0141) with variable binding affinities (Kumar et al. 2018; Oberli et al. 2014). While PFD1140w and PFE1605w showed strong binding to ATS (Kd = 0.5 μM and 5 μM, respectively), PF11_0503 and PFI1780w had moderate affinity (Kd = 500 μM and 134 μM, respectively) (Kumar et al. 2018; Oberli et al. 2014) and PFD1140w showed reduced binding to ATS upon its phosphorylation with casein kinase II, though ATS phosphorylation had no impact on its interaction with PFI1780w and PF11_0503 (Kumar et al. 2018). PFD1140w also displayed binding with MC resident *Pf* Skeleton binding protein 1 (*Pf*SBP1) in its ATS bound state, suggestive of distinct binding niches for these proteins on PFD1140w. None of the other PHISTs studied in this report bound *Pf*SBP1. Also, PFD1140w and PF11_0503 showed simultaneous interaction with ATS highlighting its ability to engage multiple PHISTs at the same time (Kumar et al. 2018). Though the functions of PFI1780w and PFD1140w are incompletely understood, these proteins have been proposed to play a role in trafficking of *Pf*EMP1 (Kumar et al. 2018; Oberli et al. 2014).

Conditional tethering of PHIST protein PFE1605w (knob-localized ATS binder) at the MCs caused no change in surface expression of *Pf*EMP1 in parasitized erythrocytes (Oberli et al. 2016). However, these parasites displayed reduced adhesion of iRBCs to the endothelial receptor CD36 highlighting the supportive role of PFE1605w in cytoadherence. Since PFE1605w depletion did not alter adhesion to host endothelial receptors other than CD36, it was hypothesised that PHIST proteins may be optimized for different *Pf*EMP1 variants (Oberli et al. 2016). In a nutshell, the *Pf*EMP1 interacting PHIST proteins from PHIST b and c subtypes display diverse functions including regulation of *var* gene expression and switching, VAR trafficking and iRBC cytoadhesion.

A PHIST family member trafficks *Pf*EMP1 and is involved in genesis of EVs

Mutant parasites lacking MC protein *PfPTP2* (*gene ID MAL7P1.172*) showed either none or severely reduced expression of PfEMP1 on the iRBC surface suggesting a role for this protein in *Pf*EMP1 trafficking (Maier et al. 2008). Later, Regev-Rudzki et al. showed localization of *Pf*PTP2 in both MCs and MC generated EVs (Regev-Rudzki et al. 2013). Disruption of its gene led to a reduced number of extracellular exosome like vesicles, suggesting a pivotal role for *Pf*PTP2 in the genesis and transmission of these structures into the extracellular space. This PHIST member has been reported to facilitate cell-cell communication amongst *P. falciparum*-infected RBCs to promote sexual differentiation of parasites (Regev-Rudzki et al. 2013).

PHIST family contributes to modification of the host cell membrane

Some of the exported parasite proteins associate with the RBC cytoskeleton and cause tremendous changes in infected erythrocyte's mechanical properties (Glenister et al. 2002; Lustigman et al. 1990). Amongst these proteins, MESA and a member of the PHISTa sub-family (*Pf*3D7_0402000) have been reported to efficiently interact with erythrocyte cytoskeleton protein 4.1r and alter its binding with other cytoskeleton proteins (Kilili and LaCount 2011; Lustigman et al. 1990; Parish et al. 2013). 4.1r is an essential mechanical linker between spectrin dimers and actin filaments, and plays a significant role in regulating the shape and mechanical stability of RBCs (Baines et al. 2014). It has a multi-domain structure

including a 30-kDa FERM domain that forms the binding site for MESA (via a 51 amino acid fragment) (Waller et al. 2003) and *Pf*3D7_0402000 (Parish et al. 2013). On the other hand, MESA binds with 4.1r through its N-terminal 13 core residue sequence (NYx[E/K]C[L/I][K/I/R][N/T] APYI**D**, where *x* is any amino acid) called the MESA erythrocyte cytoskeleton binding motif (MEC motif). Aspartic acid (**D**) is the only completely conserved amino acid within the MEC motif (Kilili and LaCount 2011).

Kilili et al. identified nine PHISTb proteins containing the MEC motif, of which four were tested and found to bind inside-out vesicles with variable affinities (Table 2) (Kilili and LaCount 2011). Also, PF10_0378 was able to immunoprecipitate band 4.1r from RBC preparations, while the others were not tested for the same. The PHISTa member '*Pf*3D7_0402000' is expressed at trophozoite and schizont stages of the parasite life cycle, co-localizes with 4.1r at the PVM (parasitophorous vacuole membrane), and interacts with 4.1r through its PHIST domain (Parish et al. 2013). Parish et al. identified an N-terminally located conserved tryptophan residue (W186) within the PHIST domain of *Pf*3D7_0402000 to be important for stringent interaction with 4.1r using the yeast-two hybrid system (Parish et al. 2013). Deletion of the extreme C-terminal region (249–325) of this protein led to complete loss of binding with 4.1r. We have analysed the sequence of this deleted C-terminal portion and found the presence of a MEC motif (266–280) within the PHIST domain (161–286) of this protein. Therefore, we propose that the MEC motif of *Pf*3D7_0402000 is likely to play a crucial role in its interaction with 4.1r. Also, binding of PF10_0378 to 4.1r and that of PHISTb proteins identified by Kilili et al. to IOVs may be driven by the MEC motif.

Multiple sequence alignment of PHIST family members and motif search has led us to identify four other PHISTb members (PFB0085c, PFA0110w, PF11_0509 and PF11_0512 with DnaJ domain) that contain a motif similar to MEC. However, no additional PHISTa or PHISTc protein seemed to carry this motif. The MEC motif is present either within or outside the PHIST domain in the identified PHIST members carrying this motif. Table 1 enlists the existing and newly identified PHIST proteins containing the MEC motif.

The binding site of 4.1r (C-lobe of FERM domain) for MESA partially overlaps with that of another RBC cytoskeletal protein 'p55' (Fig. 1a, b) (Waller et al. 2003). This causes p55 and MESA (carrying the MEC motif for binding) to compete with each other for the same binding site on 4.1r as demonstrated by competitive binding assays. Therefore, Waller et al. proposed displacement of p55 from 4.1r by MESA upon parasite infection to cause alteration of cytoskeletal protein interactions and hence membrane rigidity (Waller et al. 2003). Since the MEC motif is also found in some members of the PHIST family and other proteins (MESA like proteins), its occurrence is suggestive of convergent roles for these proteins in host cell remodelling and alteration of erythrocyte deformability (Fig. 1 c).

PHIST-b member PFE1605w (LyMP) possesses a C-terminal lysine-rich LyMP domain (321–528) apart from its N-terminal PHIST domain (159–285) (Proellocks et al. 2014). The LyMP domain is approximately 200 residues long and contains lysine-rich repeats. While the PHIST domain of PFE1605w is important for binding with the ATS domain of *Pf*EMP1, its C-terminal region is responsible for linkage with human iRBC membrane protein band 3 (Oberli et al. 2016). Band 3 is critical for maintaining the shape of uninfected erythrocytes while acting as a linker between host proteins spectrin and ankyrin (Kuhn et al. 2017). Interaction of PFE1605w with a host cytokeletal protein is likely to play a role in iRBC remodelling, though LyMP knockout parasites showed negligible alteration in membrane rigidity (Proellocks et al. 2014).

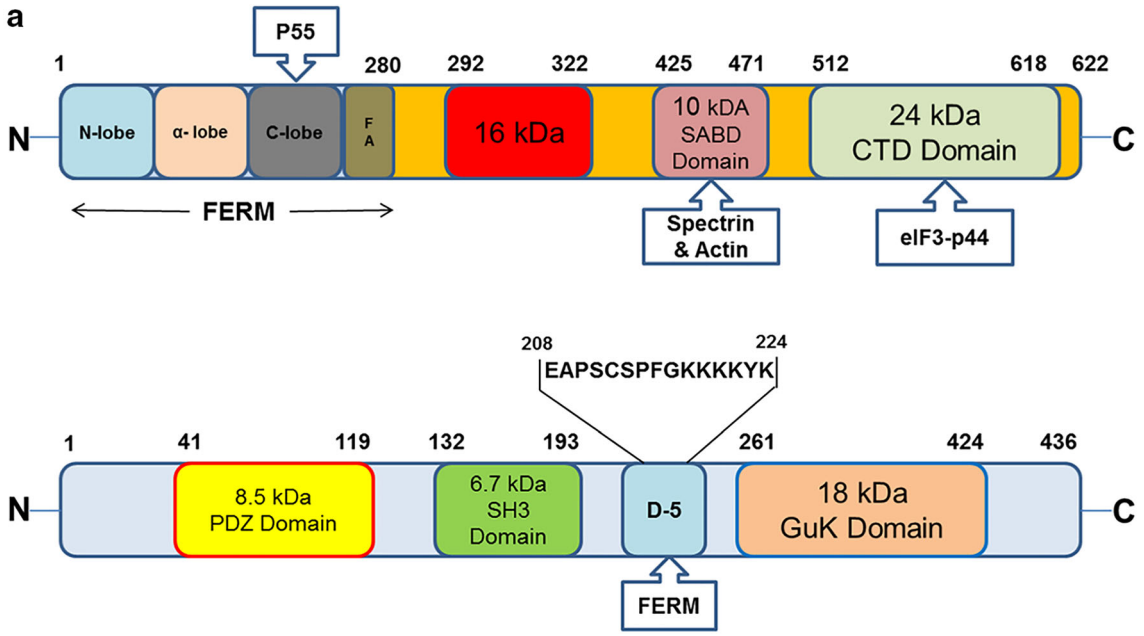
Role of PHIST proteins in sexual differentiation

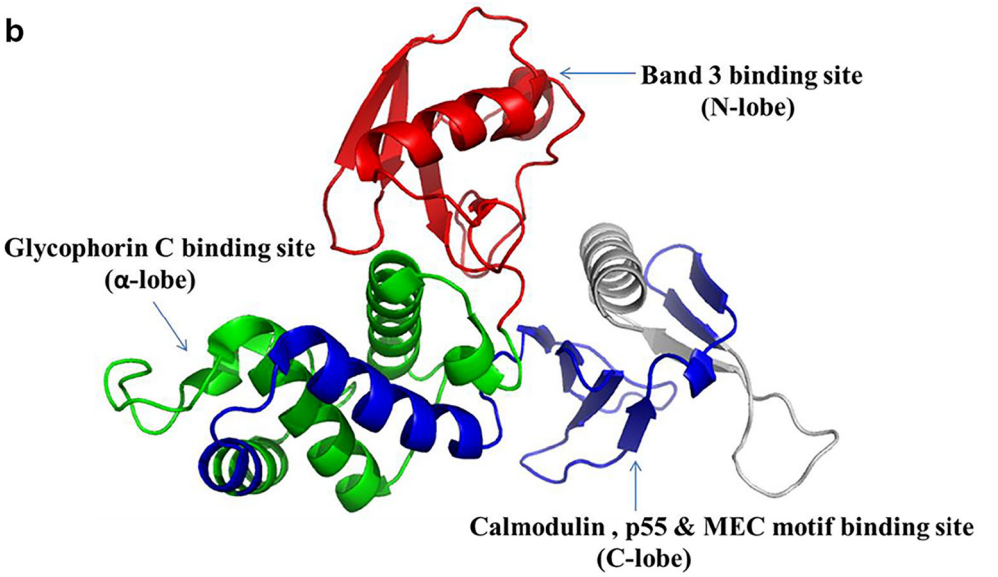
Intra-erythrocytic malaria parasites either continue to divide asexually by invasion of fresh RBCs by merozoites or produce gametocytes that are ingested by mosquitoes for commencing the sexual life cycle (Sinnis and Sim 1997). Though the underlying mechanism for this process is unclear, it has been reported that the decision between asexual multiplication or sexual differentiation occurs even before the release of merozoites from schizonts (Liu et al. 2011). The *Plasmodium* gametocyte developmental stages are critical for parasite transmission from the human host to mosquito gut, which proceed via five morphologically distinct forms. Immature *Plasmodium falciparum* gametocytes (Stages I and II) sequester with the host endothelial receptors (mostly CD36) on microvascular cells by cytoadherence via *Pf*EMP1 (Day et al. 1998; Smith et al. 2003). After stage II, *Pf*EMP1 resides inside the parasite and is not involved in their cytoadhesion. While stage III and IV gametocytes seem to bind with ICAM-1, CD49c, CD164 and CD166 on bone marrow cells, mature stage V gametocytes appear to be present in the blood stream (Rogers et al. 2000).

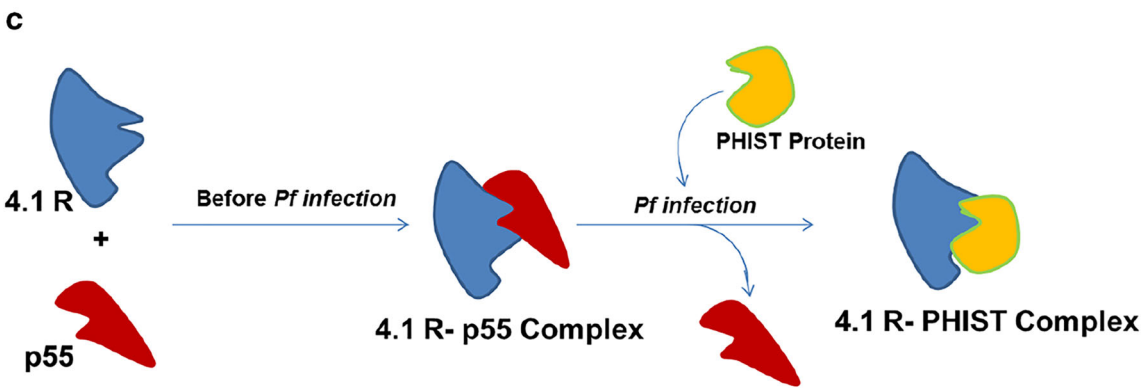
Early gametocytes express several asexual stage exported proteins of *var*, *rif* and *stevor* gene families (Petter et al. 2008). Expression of some members of the PHIST family has also been reported in the gametocytes, emphasizing their diverse roles at different developmental stages of parasite biology (Silvestrini et al. 2010). Proteomics analysis by Silvestrini et al. had identified several PHIST family proteins that are exported in early gametocytes (PHIST a-PF14_0752, PF14_0748, PF14_0744, PHIST b-PF11_0037, MAL13P1.475, MAL8P1.2, and PHIST c -PFI1770w, MAL7P1.172, PF11_0503, PFB0900c) (Silvestrini et al. 2010). Some of these were also found to express at the

Table 1 List of existing and newly identified PHIST proteins containing the MEC motif

PHIST	Previous I.D.	Sub-family	MEC-motif	Region	Function	References
PF3D7_1038800	PF10_0378	PHIST-B+DNA-J	NFCECLISAPYIDED	343–357	Bind to IOV	(Oberli et al. 2016)
PF3D7_1201100	PFL0055c	PHIST-B+DNA-J	SAVNYLECCRTARHIDKE	386–403	Bind to IOV	(Oberli et al. 2016)
PF3D7_1401600	PF14_0018	PHIST-B	PNGQYNECMKTSDYIDAQD	346–364	Bind to IOV	(Oberli et al. 2016)
PF3D7_0402100	PFD0095c	PHIST-B	NGTHYIDSIRNAPFIDLD	307–324	Bind to IOV	(Oberli et al. 2016)
PF3D7_0601500	PFF0075c	PHIST-B	SIINYKKCLIEAPYIDEY	435–453	Bind to IOV	(Oberli et al. 2016)
PF3D7_0631100	PFF1510w	PHIST-B	SIINYKKCLIEAPYIDEY	435–454	Bind to IOV	(Oberli et al. 2016)
PF3D7_0921000	PFI1030c	PHIST-B	SKHNYEAFLETVPYIDEG	319–337	Bind to IOV	(Oberli et al. 2016)
PF3D7_0937000	PFI1790w	PHIST-B	GITNYEKCLINAPYIDEE	339–357	Bind to IOV	(Oberli et al. 2016)
PF3D7_1252800	PFL2540w	PHIST-B	TIKNYTECLKMAEYVDVD	353–371	Bind to IOV	(Oberli et al. 2016)
PF3D7_0402000	PFD0090c	PHIST-A	NYITSFLEGFQNLID	266–280	Interact to 4.1 R	(MacPherson et al. 1985)
PF3D7_0201700	PFB0085c	PHIST-B+DNA-J	NYYEKMRDGLYHD	351–363	Unknown	
PF3D7_0102200	PFA0110w	PHIST-B+DNA-J	NYYDTVKDGVYLD	388–401	Unknown	
PF3D7_1149200	PF11_0509	PHIST-B+DNA-J	NYYDAVKDGKYLD	388–401	Unknown	
PF3D7_1149500	PF11_0512	PHIST-B+DNA-J	NYYDAVKDGEYLD	387–399	Unknown	


a
P55
1
280
292
322
425
471
512
618
622
N
N-lobe
α- lobe
C-lobe
F A
16 kDa
10 kDA SABD Domain
24 kDa CTD Domain
C
FERM
Spectrin & Actin
eIF3-p44
208
224
EAPSCSPFGKKKKYK
1
41
119
132
193
261
424
436
N
8.5 kDa PDZ Domain
6.7 kDa SH3 Domain
D-5
18 kDa GuK Domain
C
FERM


b
Band 3 binding site (N-lobe)
Glycophorin C binding site (α-lobe)
Calmodulin , p55 & MEC motif binding site (C-lobe)


c
4.1 R
+
p55
Before Pf infection
4.1 R- p55 Complex
PHIST Protein
Pf infection
4.1 R- PHIST Complex

◀ **Fig. 1** (**a**) Schematic representation of the domain architecture of cytoskeleton proteins 4.1R and p55. (**b**) Ribbon diagram showing 3D structure of FERM domain of 4.1 R representing binding sites for different cytoskeleton proteins. (**c**) Model representing interaction of PHIST proteins with 4.1 R

trophozoite stage. Another report described expression of PHISTa members PF14_0748 and PF14_0744 in a subpopulation of parasites that are committed for sexual differentiation, where these were found to localize in the parasitophorous vacuole of gametocytes (Eksi et al. 2005).

Role of PHIST proteins in other *Plasmodium* species

Orthologues of the PHIST protein family can be found in other *Plasmodium* species including human-infecting *P. vivax* and *P. knowlesi*, rodent-infecting *P. cynamolgi*, *P. berghei* and *P. yoelii*, and in avian *P. gallinaceum*. Interestingly, this family is best represented in *P. falciparum* (72 paralogues) followed by *P. vivax* (39 paralogues), and a poor representation in other *Plasmodium* species (Sargeant et al. 2006). Two *P. berghei* PHIST proteins PBANKA_122900 and PBANKA_114540 were found to be expressed throughout the intra-erythrocytic stages, were exported into the erythrocyte cytoplasm and localized into Maurer's cleft like vesicular structures (Moreira et al. 2016). Of these, PBANKA_114540 was shown to be essential for parasite survival at asexual stages using the gene disruption method (Moreira et al. 2016). Two other essential PHIST orthologues viz. PvPHIST/CVC-81_{95} and PcyPHIST/CVC-81_{95} of *P vivax* and *P. cynomolgi*, respectively, showed expression at asexual parasite stages and localized to the cytoplasmic side of caveolae-vesicle complexes (CVC) at the surface of the RBC membrane. CVCs are numerous specialized structures present in erythrocytes infected with *P. vivax* and *P. cynomolgi*, and resemble electron-dense knob protrusions present at the surface of *Pf* iRBCs (Akinyi et al. 2012).

PHIST proteins show specificity for *Pf*EMP1 variants

A specific *Pf*EMP1 variant expresses on knobs of an iRBC conferring it the ability to bind different host endothelial cell receptors e.g. PFD0615c, PFL0030c and PF07_0050 majorly bind CD36, CSA and ICAM-1, respectively (Smith et al. 2013). While most parasite isolates bind CD36 (Janes et al. 2011), parasite binding to ICAM-1 and CSA has been associated with cerebral and placental malaria, respectively, causing high mortality (Clausen et al. 2012; Smith et al. 2000a; Smith et al. 2000b). PfEMP1s that bind to CD36 are encoded by Var B and C genes, ICAM-1 binders by Var B genes and CSA binding PfEMP1s by Var-A subtype. Also, the adhesive property and degree of adhesion of *Pf*EMP1 significantly varies amongst different parasite isolates (Rowe et al. 2009).

It has been hypothesised that PHIST proteins may be involved with specific *Pf*EMP1 members owing to differential strength of interaction between ATS domain of PfEMP1–PHIST pairs (Oberli et al. 2016). While the PHIST 'PFE1605w' showed poor affinity for the ATS domain of PFL0030c (CSA binder), it had higher affinity for the ATS domains of PFL0020w (ICAM-1 binder) and PF08_0141 (CD36 binder). The PHIST member 'PFE1605w' is reported to play a crucial role in cytoadherence, as its conditional reduction leads to severely reduced adhesion of iRBCs to endothelial receptors [ICAM-1 (30% reduction) and CD36 (64% reduction)]. However, no effect was observed on CSA binding of these parasites (Oberli et al. 2016). This report therefore suggests that PHIST–*Pf*EMP1 pairs may have co-evolved to perform unique functions in malaria pathology.

The existence of *Pf*EMP1–PHIST pairs is also supported by transcriptome analyses. A PHISTa member 'PF14_0472' is upregulated many folds along with subsets of varA genes (HB3var3, 3D7-PFD0020c, ITvar7 and ITvar19) in cerebral malaria (Claessens et al. 2012). Some members of the PHIST family namely PFD1140w (PHISTc), PFL0050c and

Table 2 List of potential PHIST-*Pf*EMP1 partners identified using analysis of yeast two hybrid data and STRING web server

PHIST	Previous I.D.	Potential interacting *Pf*EMP1s
PF3D7_0424000	PFD1140w	PFL0030c
PF3D7_1252800	PFL2540w	PFL 1960w
PF3D7_1372000	MAL13P1.470	PFD0995c, PFL1970w, PFB0975c, PFD1235w
PF3D7_0425300	PFD1210w	PF11_0008, PF11_0007
PF3D7_1253100	PFL2555w	MAL7P1.55, PFF0010w
PF3D7_0424900	PFD1185w	MAL7P1.56, PFF1580c
PF3D7_0831750	–	PFD0995c, PFL1970w, PFB0975c, PFD1235w
PF3D7_1000700	PF10_0008	PF08_0140
PF3D7_0601700	PFF0085w	PF08_0140
PF3D7_1100600	PF11_0012	PF08_0140
PF3D7_0936800	PFI1780w	PFI1830c

```
PFI1780w/ PF3D7_0936800        -------DEI--NKCDMKKYTAEEINEMINSSNEFINRNDMNIIFSYVHESEREKFKKVE 138
PF11_0503/PF3D7_1148700        ---NNAEGYF--NNLNFNLEYTEELNEQITSLDPKTSIQELTNIFNDLNENRKKNFHYLI 139
PFD1140w/ PF3D7_0424000        NDISILEDFNNDINEALNLLSDDEIDDMINNLNDIASFEEMHNIWNELCKSEKYKFLYLI 254
PFE1605w/ PF3D7_0532400        NKVPYTSVYMGVSVQDIDYILDKKIDNL--DLYDENIIDSMKYIWNEVMDKEKKRFLGLK 202

PFI1780w/ PF3D7_0936800        ENIFKFIQSIVETYKIPDEYKMRKFKFAHFEMQGYALKQEK---------------FLLE 183
PF11_0503/ PF3D7_1148700       EALENVSLILAHQNDIPEKHRKSTWCKIKHDLQCDFIKEHP----------------NIH 183
PFD1140w/ PF3D7_0424000        YDLRKLYEELIDDIDDIEEEEEELWETCVFGVGKIQVRASGTYNDLFNN-----------303
PFE1605w/ PF3D7_0532400        NVLRKYQEEFPINNTVINNYNKKKFDKCYERIDLGADIIENGLNKMLSKCIKKNVLDKNQ 262

PFI1780w/ PF3D7_0936800        YAFLSLNGKLCERKKFKEVLEYVKREWIEFRKSMFDVWKEKLASEFREHGEMLNQKRKLK 243
PF11_0503/ PF3D7_1148700       SNFDVFIRKIQSIDEFISYIKECIHSWNNFTNEKKLQYFPKVYETYQSD-------KICK 236
PFD1140w/ PF3D7_0424000        ----LLSDEYVSKKEFIDFIKECRNRLSLIRSQLKDKCEKKIIDALS-N----------347
PFE1605w/ PF3D7_0532400        YKVLVTANRFLWRKSLLDVQEECNVILQLPTEVSKLPKKKKLRNRRKSNPNNEISLRTYG 322
```

Fig. 2 Multiple sequence alignment of the Phist domains of PFI1780w, PFD1140w, PFE1605w and PF11_0503. Residues of PFI1780w that interact with *Pf*EMP1 are marked in red

PFI1785w (both PHISTb) are reported to be transcriptionally upregulated along with *Pf*EMP-1 variant var2CSA in clinical parasite isolates (Francis et al. 2007; Ndam et al. 2008). These genes were found to be expressed many fold higher in placental parasites as compared to peripheral parasite samples from children. It has also been reported that over successive pregnancies, East African women acquire antibodies against PFD1140w which show sero-reactivity similar to that of VAR2CSA (Francis et al. 2007). Also, antisera of *P. falciparum* infected pregnant women specifically recognized recombinantly expressed PFI1785w (Ndam et al. 2008), though the role of these PHIST specific antibodies in protection against malaria is unknown. Additionally, successive pregnancies led multigravid women to acquire antibodies that inhibit binding of iRBCs to CSA causing them to be less susceptible to malaria than primagravid women (Staalsoe et al. 2001). However, the specificity of these antibodies in pregnancy associated malaria remains to be determined. Although these studies collectively indicate the role of PHIST proteins in development of cerebral and placental malaria, further studies are needed for enhanced understanding of the underlying mechanism and hence development of inhibitors/drugs against severe complicated malaria.

We have attempted to identify the PHIST–*Pf*EMP1 pairs that may operate in *P. falciparum* biology by analysing the yeast two hybrid data on Plasmodb.org (Bahl et al. 2003) and predicted a few other pairs that may interact by using the STRING web server (Szklarczyk et al. 2014). The yeast two hybrid data shows only one PHIST–*Pf*EMP1 partnership i.e. PFI1780w with PFI1830C. Known *Pf*EMP1–PHIST pairs and potential partners identified using analysis of yeast two hybrid data and STRING web server are tabulated (Table 2).

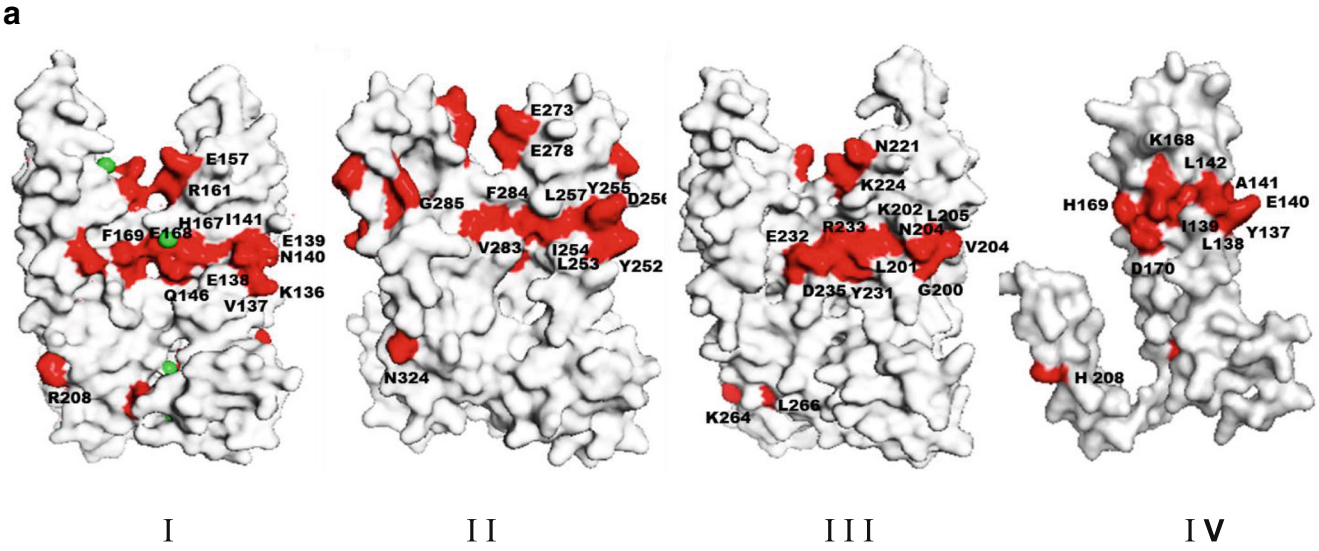


Fig. 3 Surface representation of PFI1780w (I), PFD1140w (II), PFE1605 (III) and PF11_0503 (IV). Residues of PFI1780w that interact with *Pf*EMP1 and their corresponding residues on PFD1140w (II), PFE1605 (III) and PF11_0503 (IV) are marked in red

The table shows different PHIST members that seem to pair with *Pf*EMP1s having CD36, CD28, ICAM-1 and CSA binding specificities. Half of the predicted PHIST members shown in the table were predicted to be pseudogenes. However, of these, expression of peptides specific for PF3D7_0601700/PFF0085w was seen in merozoites according to mass spectrometric data present on Plasmodb.org (Bahl et al. 2003).

Structure analysis of PHIST-ATS complex

Oberli et al. solved the first crystallographic structure of the PHIST domain of PFI1780w (PDB ID. 4JLE) and described a partial model for the PHIST-ATS complex using nuclear magnetic resonance (NMR) (Oberli et al. 2014). The atomic resolution structure of PHIST domain of PFI1780w shows the protein to be dimeric, comprising of three long α-helices separated from a short α-helix by a flexible loop. This study also identified the ATS-binding residues to be located on α2 and α3 helices of the PHIST domain of PFI1780w. These include amino acid stretches 'KVEENI' (136 to 141) of α2 and 'HFExQ' (167 to 171) of α3 helices of the PHIST domain. While the PHIST domains of PFI1780w and PFD1140w were seen to exist as dimers, PF11_0503 was present in its monomeric state (Kumar et al. 2018; Oberli et al. 2014). Kumar et al. compared homology models of PFD1140w and PF11_503 with the crystal structure of PFI1780w to identify several dimer interface residues that were present in the dimeric forms but went missing in PF11_0503, causing it to remain as a monomer (Kumar et al. 2018).

Since PHIST proteins showed differential binding affinity to ATS domains of *Pf*EMP1s (Oberli et al. 2016), we compared the residues of PFI1780w involved in ATS binding with those in PFD1140w, PF11_0503 and PFE1605w. Multiple sequence alignment of the PHIST

```
PF3D7_0800200      NTDVWIEIDMDDPKGKKEFSNMDT------------IL----------------------      2796
PF3D7_1300300      NTDVSIEIDMDDPKGKKEFSNMDT------------IL----------------------      3285
PF3D7_1100200      NTDVSIQIDIDENKGKKEFSNMDT------------NVDTPTMDS-ILD--DLET-YNEP      2931
PF3D7_0632500      NTDVSIQIDMNNPKTTNEFTYVDSNPNQVDDTYVDSNPDNSSMDT-ILE--DLDKPFNEP      3898
PF3D7_0600200      NTDVSIQIHMDNPKPINQFT----------------------NMDT-ILE--DLEK-YNEP      2825
PF3D7_1200400      NTDVSIQIHMDNPKPINQFT----------------------NMDT-ILE--DLEK-YNEP      2806
PF3D7_0712400      NTDVSIQIHMDDPKPINQFT----------------------NMDT-ILE--DLEK-YNEP      2010
PF3D7_1240300      NTDVSIQIHMDNPKPINEFTYVDSNPNQVDDTYVDSNPDNSSMDT-ILE--DLEK-YNEP      2590
PF3D7_0632800      NTDVSIQIHMDNPKPINQFT----------------------NMDT-ILD--DLDKPFNEP      2181
PF3D7_0200100      NTDVSIQIHMDNPKPINQFT----------------------NMDT-ILE--DLEK-YNEP      1657
PF3D7_0324900      NTDVSIQIDMDNPKPINQFT----------------------NMDT-ILE--DLEK-YNEP      2115
PF3D7_1200600      RTNISMDTYIDET-----------------------NNNNVVATSIIGDDQMENSYNC-      3056
PF3D7_1150400      NTDVSIQIDMDDGKPKKEFSNMDT-----------------ILDD-M-------------      3061
PF3D7_0617400      NSDVSIQIDMNSKPI--------------------------------------------      2394
PF3D7_0400400      NTDVSIQIDMDENKGKKEFSNMDT------------NVDTPTMDN-ILD--DLET-CNEP      3404
PF3D7_0425800      NTDVSIQIDMDETKGKKEFSNMDT-----------------ILDD-M-------------      3495
PF3D7_0100300      NTDVSIQIDIDENKGKKDLSNMDT-----------------ILDD-M-------------      1262
PF3D7_0937600      NTDVSIEIDMDNPKPINQFSNMDI------------NVDTPTMDN-M-------------      1250
PF3D7_0600400      NTDVSIEIDMDNPKPINQFSNMDI------------NVDTPTMDN-M-------------      1260
PF3D7_0809100      NTDVSIQIDMHNPKPKNEFTNMDTNP------------DNFIKDT-ILN--DLEK-HREP      2052
PF3D7_1200100      NTDVSIQIDMNNPKTTNEFTYVDSNPNQVDDTYVDSNPDNSSMDT-ILE--DLDKPFNEP      2123
PF3D7_0800100      NTDVSIQIDMDHEKRMKEFT----------------------NMDT-ILE--DLDKPFNEP      2064
PF3D7_0426000      NTDVSIQIDMNNPKTTNEFTYVDSNPNQVDDTYVDSNPDNSSMDT-ILE--DLDKPFNEP      2093
PF3D7_1255200      NTDVSIQIDMDNPKPTNEFTYVDSNPNQVDDTYVDSNPDNSSMDT-ILD--DLE-KYNEP      2212
PF3D7_0400100      NTDVSIQIDMDNPKPINQFTN----------------------MDI-NVD--TP------T      2592
PF3D7_0115700      NTDVSIQIHMDNPKPINEFTN----------------------MDT-ILD--DLK-TYNEP      2149
PF3D7_0100100      NTDVSIQIDMDHEKRMKEFTN----------------------MDT-ILD--DLK-TYNEP      2109
PF3D7_0800300      NTDVSIQIDMDNPKTTNEFTYVDSNPNQVDDTYVDSNPDNSSMDT-ILE--DLEK-YKEP      2926
PF3D7_0712800      NTDVSIQIHMDNPKPKNEFKNMDTT------------PNKSTMDT-MLD--DLEK-YNEP      2144
PF3D7_1041300      NTDVSIQIHMDNPKPINQFTN----------------------MDT-ILE--DLDK-YNEP      2140
PF3D7_0300100      NTDVSIQIDMDDGKPKKEFTN----------------------MDT-ILE--DLDK-YNEP      2155
PF3D7_0500100      NTDVSIQIDMDDGKPKKEFTN----------------------MDT-ILE--DLDK-YNEP      2153
PF3D7_1300100      NTDVSIQIDMNNPKTTNEFTYVDSNPNQVDDTYVDSNPDNSSMDT-ILE--DLDKPFNEP      2105
PF3D7_0412400      NSDVSIQIHMDNPKPTNEDNVVDCNPVG-NNIYVDNNPNQTFPSN---------------      2169
PF3D7_0700100      NTDVSIQIHMDNPKTTNEFTYVDSNPNQVDDTYVDSNPDNSSMDT-ILE--DLDK-YNEP      2232
PF3D7_1373500      NTDVSIQIDMNNPKTTNEFTYVDSKPNQVDDTYVDSNPDNSSMDT-ILD--DLDKPFNEP      2167
PF3D7_0712300      NSDVSIQIDMDNPNQVDDNTYLDT------------YPDKSTMDT-IMD--DLEK-YNEP      2216
PF3D7_0900100      NTDVSIQIHMDNPKPINQFTNMDT------------ILEDLDKYN-------------EP      2216
```

Fig. 4 Multiple sequence alignment of the intracellular acidic terminal segment (ATS domain) of *Pf*EMP1s. Yellow colour represents PHIST binding residues of PF3D7_0800200/PF08_0141 and their corresponding residues in other *Pf*EMP1s

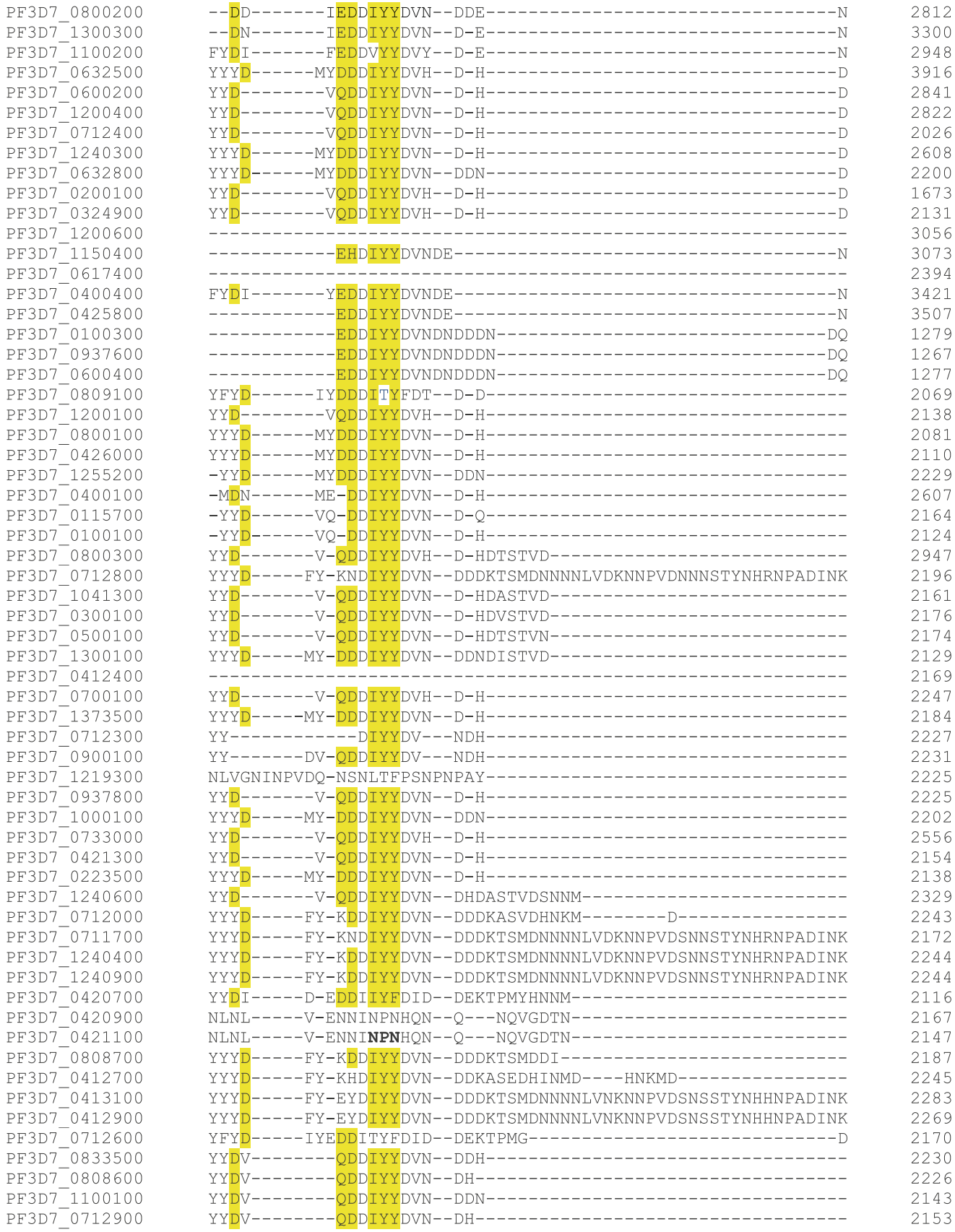


Fig. 4 (continued)

```
PF3D7_0800200      PSVDNIPMDHNKV--DVPKKVHVEMKILNNTSN---GSLEPEFPISDVWNI      2858
PF3D7_1300300      PSVNDIPMDHNKV--DVPKKVHVEMKILNNTST---GSLEQQFPISDVWNI      3346
PF3D7_1100200      PFVDDIPMDHNKV--DVPKKVHIEMKILNNTSN---GSLEQQFPISDVWNI      2994
PF3D7_0632500      T----STVDTNAM--DVPSKVQIEMDVN----T---KLVKEKYPISDVWDI      3954
PF3D7_0600200      A----STVDSNAV--NVPSKVQIEMDVN----T---KLVKEKYPIADVWDI      2879
PF3D7_1200400      A----STVDSNNM--DVPSKVQIEMDVN----T---KLVKEKYPIADVWDI      2860
PF3D7_0712400      T----STADSNAM--DVPSKVQIEMDVN----T---KLVKEKYPIADVWDI      2064
PF3D7_1240300      A----STVDSNNM--DVPSKVQIEMDVN----T---KLVKEKYPIADVWDI      2646
PF3D7_0632800      I----STVDTNAM--DVPSKVQIEMDVN----T---KLVKEKYPIADVWDI      2238
PF3D7_0200100      V----STAGSNAM--DVPSKVQIEMDIN----T---KLVKEKYPISDVWDI      1711
PF3D7_0324900      T----STVDSNNM--DVPSKVQIEMDVN----T---KLVKEKYPIGDVWDI      2169
PF3D7_1200600      --------------------------------------------------      3056
PF3D7_1150400      PSVNDIPMDHNKV--DVPKKVHVEMKILNNTSN---GSLEPEFPISDVWNI      3119
PF3D7_0617400      --------------------------------------------------      2394
PF3D7_0400400      PSVDDIPMDHNKV--DVPKKVHVEMKILNNTSN---GSLEPEFPISDVWNI      3467
PF3D7_0425800      PSVDDIPMDHNKV--DVPKKVHVEMKILNNTSN---GSLEQQFPISDVWNI      3553
PF3D7_0100300      PSVYDISMDHNKVDVDVPKKVHIEMKILNNTSN---GSLEQQFPISDVWNI      1327
PF3D7_0937600      PSVYDIPMDHNKVDVDVPKKVHIEMKILNNTSN---GSLEQQFPISDVWNI      1315
PF3D7_0600400      PSVYDIPMDHNKVDVDVPKKVHIEMKILNNTSN---GSLEQQFPISDVWNI      1325
PF3D7_0809100      ---VKPPMDD------IHIKEQTEMNALHNNKMN--ELLEKEYPISDIWNI      2109
PF3D7_1200100      ---DTSTVDTNAM--DVPSKVQIEMS----VKNH--KLVKEKYPIGDVWDI      2178
PF3D7_0800100      ---DTSTVDSNAM--DVPSKVQIEMD----VN-T--KLVKEKYPIGDVWDI      2120
PF3D7_0426000      ---DASTVDSNNM--DVPSKVQIEMD----I-NT--KLVKEKYPIGDVWDI      2149
PF3D7_1255200      ---DISTVDTNAM--DVPSKVQIEMD----V-NT--KLVKEKYPISDVWDI      2268
PF3D7_0400100      ---DTSTVDSNTM--DVPSKVQIEMD----V-NT--KLVKEKYPIADVWDI      2646
PF3D7_0115700      ---DVSTVDSNAM--DVPSKVQIEMD----V-NT--KLVKEKYPIGDVWDI      2203
PF3D7_0100100      ---DASTVDSNNM--DVPSRVQIEMD----V-NT--KLVKEKYPIADVWDI      2163
PF3D7_0800300      ---------TNAM--DVPSKVQIEMD----INT---KLVKEKYPIGDVWDI      2980
PF3D7_0712800      NFVHKNNQNQHPI--EKPTKIQIEMN----SNNR--EVVEQQYPIADIWNI      2239
PF3D7_1041300      ---------SNTM--DIPSKVQIEMD----VNT---KLVKEKYPIADVWDI      2194
PF3D7_0300100      ---------SNAM--DIPSKVQIEMD----VNT---KLVKEKYPIADVWDI      2209
PF3D7_0500100      ---------PNNM--EKPSKVKIELD----VNK---KTIKEKYPISDVWDI      2207
PF3D7_1300100      ---------TNAM--DVPSKVQIEMD----VNT---KLVKEKYPISDVWDI      2162
PF3D7_0412400      ---PNPVENNTYV--NAPTNVQIEMD----VNNHK--VVKEKYPISDMLDI      2209
PF3D7_0700100      ---DTSTVDSNAM--DVPSKVQIEMD----VN-TK--LVKEKYPIADVWDI      2286
PF3D7_1373500      ---DASTVDTNAM--DVPSKVQIEMD----VN-TK--LVKEKYPIADVWDI      2223
PF3D7_0712300      ---DASNVDSNNM--DVPSKVKIEMS----VK-NT-QMMEGKYPIGDVWDI      2267
PF3D7_0900100      ---DTSTVDTNAM--DVPSKVQIEMD----VN-TK--LVKEKYPIADVWDI      2270
PF3D7_1219300      ---DNIYIDHNNE--DLPSKVQIEMS----VK-NG--EMAKEK--------      2256
PF3D7_0937800      ---DASVVDSNAV--DVPSKVQIEMD----IN-TK--LVKEKYPITDVWDI      2264
PF3D7_1000100      ---DTSTVDSNNM--DVPSKVQIEMD----VN-TK--LVKEKYPISDVWDI      2241
PF3D7_0733000      ---DASTVDSNAM--DVPSKVQIEMD----VN-TK--LVKEKYPISDVWDI      2595
PF3D7_0421300      ---DTSTVDTNAM--DVPSKVQIEMD----VN-TK--LVKEKYPIGDVWDI      2193
PF3D7_0223500      ---DTSTVDTNAM--DVPSKVQIEMD----VN-TK--LVKEKYPIADVWDI      2177
PF3D7_1240600      ---------------DVPSKVKLEMS----V-KNT-QMMEEKYPIGDVWDI      2359
PF3D7_0712000      ---------NNNS--DVPTKVQIEMN----VINNQ-ELLQNEYPISHM---      2275
PF3D7_0711700      TFVDINNHNQHPI--EKPTKIQIEMN----S-NNR-EVVEQQYPIADIWNI      2215
PF3D7_1240400      NFVDKNNQNQHPI--EKPTKIQIEMN----S-NNR-EVVEQQYPIADIWNI      2287
PF3D7_1240900      NFVDKNNQNQHPI--EKPTKIQIEMN----I-NNG-ELVKEKYPISDIWNI      2287
PF3D7_0420700      --------DNNKS--NVPTKVQIEMN----VI-NKQELFQEEFPISDIWNI      2152
PF3D7_0420900      --------FVDTP--TNPTNVQIEMD----VN---TKLVKEKYPIGDVWDI      2201
PF3D7_0421100      --------FVDTP--TNPTNVQIEMD----VN---TKLVKEKYPIGDVWDI      2181
PF3D7_0808700      -YVDHNNVTSNNM--DVPTKMHIEMN----IVNNKNEIFEEEYPISDIWNI      2231
PF3D7_0412700      ---------NNNS--DVPTNVQIEMN----VINNQ-ELLQNEYPISHM---      2277
PF3D7_0413100      TFVDINNHNQHPI--EKPTKIQIEMN----S-NNR-EVDEQQYPIADIWNI      2326
PF3D7_0412900      TFVDINNHNQHPI--EKPTKIQIEMN----S-NNR-EVDEQQYPIADIWNI      2312
PF3D7_0712600      IYVDHNNVNSNNM--DVPNKVHIEMN----IVNNKKEIFEEEYPISDIWNI      2215
PF3D7_0833500      ---DTSTVDSNNM--DVPSKVQIEMD----VN---TKLVKEKYPIADVWDI      2269
PF3D7_0808600      ---DTSTVDSNAM--DIPSKVQIEMD----VN---TKLVKEKYPIGDVWDI      2265
PF3D7_1100100      ---DISTVDSNNM--DIPSKVQIEMD----VN---TKLVKEKYPIADVWDI      2182
PF3D7_0712900      ---DASTVDSNAM--DVPSKVQIEMD----VN---TKLVKEKYPIADVWDI      2192
```

Fig. 4 (continued)

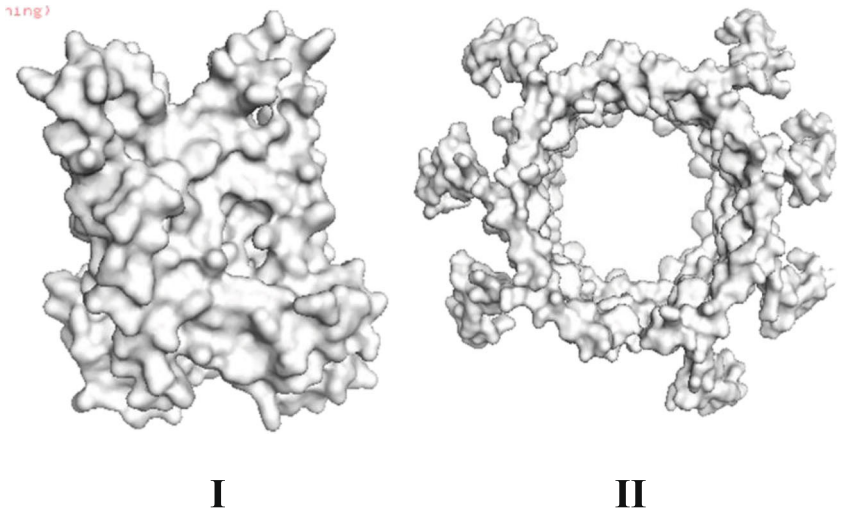


Fig. 5 Surface representation of modelled structures of PF10_0378 (I) and PF10_0378 (II)

domains of these proteins showed poor conservation of the ATS binding amino acids (Fig. 2). Residue mapping on their modelled structures showed most of the ATS interacting residues to lie on the protein surface (Fig. 3) (18, homology model of PFE1605w generated using Swiss model (Guex et al. 2009), and validated using Rampage (Lovell et al. 2003) and Errat (Colovos and Yeates 1993). An attempt to identify the PHIST binding niche on the ATS domain (PF08_0141) revealed that its PHIST binding region was a part of disordered stretches present C-terminal to the core ATS domain. However, multiple sequence alignment showed PHIST binding residues on most ATS domains to be highly conserved (Fig. 4). Existence of gross sequence diversity amongst PHIST members and lack of conservation in their ATS binding residues highlights the expected functional diversity in the PHIST family.

Since the PHIST domains of different PHIST proteins show significant sequence variability, we attempted to predict reliable homology models for the PHIST domains of different members of this family using SWISS-MODEL (Guex et al. 2009) (Table S2). While most PHIST domains chose the solved crystal structure of PFI1780w (PDB ID: 4JLE) as the template for model building, three PHIST members (PF10_0378, PF14_0748 and PF14_0745) identified templates other than 4JLE owing to poor sequence identity and/or low query cover. PF10_0378 used the J domain of Sis1 protein, a hsp40 co-chaperone from *Saccharomyces cereviseae* (4RWU) as a template for homology modelling, whereas PF14_0748 identified bovine endothelial nitric oxide synthase heme domain (4UPQ) (Fig. 5). Interestingly, the model corresponding to PF10_0378 was all helical carrying an overall fold very similar to PFI1780w suggestive of structural convergence. The overall architecture and assembly of the model generated for PF14_0748 was starkly different from 4JLE despite displaying an all helical structure (Fig. 5). The model was obtained for a portion of the PHIST domain only, which encompasses the N-terminal ~60 amino acid TIGR01639 domain whose functions are unknown.

Conclusion

PHIST comprises one of the significant exported protein families of malaria parasites. PHIST family proteins are unique to *Plasmodium* and have linearly expanded to give a large variation in the number of members present in different *Plasmodium* species. *Plasmodium falciparum* carries the highest number of PHIST proteins. However, it remains to be seen whether there is any correlation between virulence and the number of PHIST proteins in various *Plasmodium* species.

The PHIST family proteins localize to different subcellular regions in *Plasmodia*-infected erythrocytes. PHIST proteins are not only widely distributed in different asexual parasitic stages but their expression has also been observed in early gametocytes that are committed for sexual differentiation. PHIST proteins play several key roles including *Pf*EMP1 trafficking, regulation of its expression and *var* gene switching along with controlling iRBC cytoadhesion. Several PHIST members have been found in extracellular vesicles that have key effects on malaria pathogenesis, with one of these participating in EV generation. Nearly all the studied PHIST proteins have shown direct interaction with the ATS domain of *Pf*EMP1, while another PHIST member displayed binding with the extracellular domain of *Pf*EMP1. Based on varying binding affinities of PHIST proteins with an ATS domain of *Pf*EMP1, their export pattern and localization at iRBCs, it is believed that PHIST proteins may be optimized for specific *Pf*EMP1 variants. Members of the PHIST family also control gametocytogenesis and modification of the host cell. Domains other than PHIST also seem to play significant functions in parasite biology including the LyMP domain present in a

single PHIST member and the more widespread MEC-motif that contributes to host cell remodelling. Only a few members of this functionally diverse family have been explored till date highlighting the continued need for a better understanding of the PHIST family. Also, considering their important roles in malaria biology, these members seem to be attractive candidates for future anti-malarial drug development. Since a significant proportion of PHIST family proteins seem to be essential for parasite survival, they form attractive drug and vaccine candidates in the fight against *Plasmodium falciparum*-caused malaria.

References

Abdi A, Yu L, Goulding D, Rono MK, Bejon P, Choudhary J, Rayner J (2017) Proteomic analysis of extracellular vesicles from a Plasmodium falciparum Kenyan clinical isolate defines a core parasite secretome. Wellcome open research 2:50

Akinyi S, Hanssen E, Meyer EV, Jiang J, Korir CC, Singh B, Lapp S, Barnwell JW, Tilley L, Galinski MR (2012) A 95 kDa protein of Plasmodium vivax and P. cynomolgi visualized by three-dimensional tomography in the caveola–vesicle complexes (Schüffner's dots) of infected erythrocytes is a member of the PHIST family. Mol Microbiol 84:816–831

Bahl A, Brunk B, Crabtree J, Fraunholz MJ, Gajria B, Grant GR, Ginsburg H, Gupta D, Kissinger JC, Labo P, Li L (2003) PlasmoDB: the Plasmodium genome resource A database integrating experimental and computational data. Nucleic Acids Res 31: 212–215

Baines AJ, Lu HC, Bennett PM (2014) The protein 4.1 family: hub proteins in animals for organizing membrane proteins. Biochimica et Biophysica Acta (BBA)-Biomembranes 1838:605–619

Claessens A, Adams Y, Ghumra A, Lindergard G, Buchan CC, Andisi C, Bull PC, Mok S, Gupta AP, Wang CW, Turner L (2012) A subset of group A-like var genes encodes the malaria parasite ligands for binding to human brain endothelial cells. Proc Natl Acad Sci 109: E1772–E1781

Clausen TM, Christoffersen S, Dahlbäck M, Langkilde AE, Jensen KE, Resende M, Agerbæk MØ, Andersen D, Berisha B, Ditlev SB, Pinto VV (2012) Structural and functional insight into how the Plasmodium falciparum VAR2CSA protein mediates binding to chondroitin sulfate a in placental malaria. J Biol Chem 287: 23332–23345

Colovos C, Yeates TO (1993) Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci 2:1511–1519

Crabb BS, Cooke BM, Reeder JC, Waller RF, Caruana SR, Davern KM, Wickham ME, Brown GV, Coppel RL, Cowman AF (1997) Targeted gene disruption shows that knobs enable malaria-infected red cells to cytoadhere under physiological shear stress. Cell 89: 287–296

Day KP, Hayward RE, Smith D, Culvenor JG (1998) CD36-dependent adhesion and knob expression of the transmission stages of Plasmodium falciparum is stage-specific. Mol Biochem Parasitol 93:167–177

Eksi S, Haile Y, Furuya T, Ma L, Su X, Williamson KC (2005) Identification of a subtelomeric gene family expressed during the asexual–sexual stage transition in Plasmodium falciparum. Mol Biochem Parasitol 143:90–99

Francis SE, Malkov VA, Oleinikov AV, Rossnagle E, Wendler JP, Mutabingwa TK, Fried M, Duffy PE (2007) Six genes are preferentially transcribed by the circulating and sequestered forms of Plasmodium falciparum parasites that infect pregnant women. Infect Immun 75:4838–4850

Glenister FK, Coppel RL, Cowman AF, Mohandas N, Cooke BM (2002) Contribution of parasite proteins to altered mechanical properties of malaria-infected red blood cells. Blood 99:1060–1063

Goel S, Muthusamy A, Miao J, Cui L, Salanti A, Winzeler EA, Gowda DC (2014) Targeted disruption of a ring-infected erythrocyte surface antigen (RESA)-like export protein gene in Plasmodium falciparum confers stable chondroitin 4-sulfate cytoadherence capacity. J Biol Chem 289:34408–34421

Guex N, Peitsch MC, Schwede T (2009) Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. Electrophoresis 30:S162–S173

Haldar K, Mohandas N (2007) Erythrocyte remodeling by malaria parasites. Curr Opin Hematol 14:203–209

Hiller NL, Bhattacharjee S, van Ooij C, Liolios K, Harrison T, Lopez-Estrano C, Haldar K (2004) A host-targeting signal in virulence proteins reveals a secretome in malarial infection. Science 306: 1934–1937

Janes JH, Wang CP, Levin-Edens E, Vigan-Womas I, Guillotte M, Melcher M, Mercereau-Puijalon O, Smith JD (2011) Investigating the host binding signature on the Plasmodium falciparum PfEMP1 protein family. PLoS Pathog 7:e1002032

Kilili GK, LaCount DJ (2011) An erythrocyte cytoskeleton-binding motif in exported Plasmodium falciparum proteins. Eukaryot Cell 10: 1439–1447

Kirk K, Saliba KJ (2007) Targeting nutrient uptake mechanisms in Plasmodium. Curr Drug Targets 8:75–88

Kuhn V, Diederich L, Keller TS IV, Kramer CM, Lückstädt W, Panknin C, Suvorava T, Isakson BE, Kelm M, Cortese-Krott MM (2017) Red blood cell function and dysfunction: redox regulation, nitric oxide metabolism, anemia. Antioxid Redox Signal 26:718–742

Kumar V, Kaur J, Singh AP, Singh V, Bisht A, Panda JJ, Mishra PC, Hora R (2018) PHISTc protein family members localize to different subcellular organelles and bind Plasmodium falciparum major virulence factor PfEMP-1. FEBS J 285:294–312

Lanzer M, Wickert H, Krohne G, Vincensini L, Breton CB (2006) Maurer's clefts: a novel multi-functional organelle in the cytoplasm of Plasmodium falciparum-infected erythrocytes. Int J Parasitol 36: 23–36

Liu Z, Miao J, Cui L (2011) Gametocytogenesis in malaria parasite: commitment, development and regulation. Future Microbiol 6: 1351–1369

Lovell SC, Davis IW, Arendall WB III, De Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC (2003) Structure validation by Cα geometry: ϕ, ψ and Cβ deviation. Proteins: Structure Function Bioinformatics 50:437–450

Lustigman S, Anders RF, Brown GV, Coppel RL (1990) The mature-parasite-infected erythrocyte surface antigen (MESA) of Plasmodium falciparum associates with the erythrocyte membrane skeletal protein, band 4.1. Mol Biochem Parasitol 38:261–270

MacPherson GG, Warrell MJ, White NJ, Looareesuwan SO, Warrell DA (1985) Human cerebral malaria. A quantitative ultrastructural analysis of parasitized erythrocyte sequestration. Am J Pathol 119:385

Maier AG, Cooke BM, Cowman AF, Tilley L (2009) Malaria parasite proteins that remodel the host erythrocyte. Nat Rev Microbiol 7: 341–354

Maier AG, Rug M, O'Neill MT, Brown M, Chakravorty S, Szestak T, Chesson J, Wu Y, Hughes K, Coppel RL, Newbold C (2008) Exported proteins required for virulence and rigidity of Plasmodium falciparum-infected human erythrocytes. Cell 134: 48–61

Marti M, Good RT, Rug M, Knuepfer E, Cowman AF (2004) Targeting malaria virulence and remodeling proteins to the host erythrocyte. Science 306:1930–1933

Moreira CK, Naissant B, Coppi A, Bennett BL, Aime E, Franke-Fayard B, Janse CJ, Coppens I, Sinnis P, Templeton TJ (2016) The Plasmodium PHIST and RESA-like protein families of human and rodent malaria parasites. PLoS One 11:e0152510

Ndam NT, Bischoff E, Proux C, Lavstsen T, Salanti A, Guitard J, Nielsen MA, Coppée JY, Gaye A, Theander T, David PH (2008) Plasmodium falciparum transcriptome analysis reveals pregnancy malaria associated gene expression. PLoS One 3:e1855

Oberli A, Slater LM, Cutts E, Brand F, Mundwiler-Pachlatko E, Rusch S, Masik MF, Erat MC, Beck HP, Vakonakis I (2014) A Plasmodium falciparum PHIST protein binds the virulence factor PfEMP1 and comigrates to knobs on the host cell surface. FASEB J 28:4420–4433

Oberli A, Zurbrügg L, Rusch S, Brand F, Butler ME, Day JL, Cutts EE, Lavstsen T, Vakonakis I, Beck HP (2016) Plasmodium falciparum Plasmodium helical interspersed subtelomeric proteins contribute to cytoadherence and anchor P. falciparum erythrocyte membrane protein 1 to the host cell cytoskeleton. Cell Microbiol 18:1415–1428

Parish LA, Mai DW, Jones ML, Kitson EL, Rayner JC (2013) A member of the Plasmodium falciparum PHIST family binds to the erythrocyte cytoskeleton component band 4.1. Malar J 12:160

Petter M, Bonow I, Klinkert MQ (2008) Diverse expression patterns of subgroups of the rif multigene family during Plasmodium falciparum gametocytogenesis. PLoS One 3:e3779

Pongponratn E, Riganti M, Punpoowong B, Aikawa M (1991) Microvascular sequestration of parasitized erythrocytes in human falciparum malaria: a pathological study. Am J Trop Med Hyg 44:168–175

Prajapati SK, Singh OP (2013) Remodeling of human red cells infected with Plasmodium falciparum and the impact of PHIST proteins. Blood Cell Mol Dis 51:195–202

Proellocks NI, Herrmann S, Buckingham DW, Hanssen E, Hodges EK, Elsworth B, Morahan BJ, Coppel RL, Cooke BM (2014) A lysine-rich membrane-associated PHISTb protein involved in alteration of the cytoadhesive properties of Plasmodium falciparum-infected red blood cells. FASEB J 28:3103–3113

Regev-Rudzki N, Wilson DW, Carvalho TG, Sisquella X, Coleman BM, Rug M, Bursac D, Angrisano F, Gee M, Hill AF, Baum J (2013) Cell-cell communication between malaria-infected red blood cells via exosome-like vesicles. Cell 153:1120–1133

Rogers NJ, Hall BS, Obiero J, Targett GA, Sutherland CJ (2000) A model for sequestration of the transmission stages of Plasmodium falciparum: adhesion of gametocyte-infected erythrocytes to human bone marrow cells. Infect Immun 68:3455–3462

Rowe JA, Claessens A, Corrigan RA, Arman M (2009) Adhesion of Plasmodium falciparum-infected erythrocytes to human cells: molecular mechanisms and therapeutic implications. Expert Rev Mol Med 11:e16

Sanderson T, Rayner JC (2017) PhenoPlasm: a database of disruption phenotypes for malaria parasite genes. Wellcome Open Res 2:45

Sargeant TJ, Marti M, Caler E, Carlton JM, Simpson K, Speed TP, Cowman AF (2006) Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. Genome Biol 7:R12

Sharma L, Shukla G (2017) Placental malaria: a new insight into the pathophysiology. Frontiers in medicine 4:117

Sherman IW, Eda S, Winograd E (2003) Cytoadherence and sequestration in Plasmodium falciparum: defining the ties that bind. Microbes Infect 5:897–909

Silvestrini F, Lasonder E, Olivieri A, Camarda G, van Schaijk B, Sanchez M, Younis SY, Sauerwein R, Alano P (2010) Protein export marks the early phase of gametocytogenesis of the human malaria parasite Plasmodium falciparum. Mol Cell Proteomics 9:1437–1448

Sinnis P, Sim BK (1997) Cell invasion by the vertebrate stages of Plasmodium. Trends Microbiol 5:52–58

Smith JD, Craig AG, Kriek N, Hudson-Taylor D, Kyes S, Fagen T, Pinches R, Baruch DI, Newbold CI, Miller LH (2000a) Identification of a Plasmodium falciparum intercellular adhesion molecule-1 binding domain: a parasite adhesion trait implicated in cerebral malaria. Proc Natl Acad Sci 97:1766–1771

Smith JD, Rowe JA, Higgins MK, Lavstsen T (2013) Malaria's deadly grip: cytoadhesion of P lasmodium falciparum-infected erythrocytes. Cell Microbiol 15:1976–1983

Smith JD, Subramanian G, Gamain B, Baruch DI, Miller LH (2000b) Classification of adhesive domains in the Plasmodium falciparum erythrocyte membrane protein 1 family. Mol Biochem Parasitol 110:293–310

Smith TG, Serghides L, Patel SN, Febbraio M, Silverstein RL, Kain KC (2003) CD36-mediated nonopsonic phagocytosis of erythrocytes infected with stage I and IIA gametocytes of Plasmodium falciparum. Infect Immun 71:393–400

Soni R, Sharma D, Bhatt TK (2016) Plasmodium falciparum secretome in erythrocyte and beyond. Front Microbiol 7:194

Staalsoe T, Megnekou R, Fievét N, Ricke CH, Zornig HD, Leke R, Taylor DW, Deloron P, Hviid L (2001) Acquisition and decay of antibodies to pregnancy-associated variant antigens on the surface of Plasmodium falciparum-infected erythrocytes that protect against placental parasitemia. J Infect Dis 184:618–626

Su XZ, Heatwole VM, Wertheimer SP, Guinet F, Herrfeldt JA, Peterson DS, Ravetch JA, Wellems TE (1995) The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of Plasmodium falciparum-infected erythrocytes. Cell 82:89–100

Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M (2014) STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic Acids Res 43:D447–D452

Tarr SJ, Moon RW, Hardege I, Osborne AR (2014) A conserved domain targets exported PHISTb family proteins to the periphery of Plasmodium infected erythrocytes. Mol Biochem Parasitol 196:29–40

Waller KL, Nunomura W, An X, Cooke BM, Mohandas N, Coppel RL (2003) Mature parasite-infected erythrocyte surface antigen (MESA) of Plasmodium falciparum binds to the 30-kDa domain of protein 4.1 in malaria-infected red blood cells. Blood 102:1911–1914

Warncke JD, Vakonakis I, Beck HP (2016) Plasmodium helical interspersed subtelomeric (PHIST) proteins, at the center of host cell remodeling. Microbiol Mol Biol Rev 80:905–927

World Health Organization (2016) World malaria report 2015. World Health Organization, Geneva

Zhang M, Wang C, Otto TD, Oberstaller J, Liao X, Adapa SR, Udenze K, Bronner IF, Casandra D, Mayho M, Brown J (2018) Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis. Science 360:eaap7847

Zhang Q, Ma C, Oberli A, Zinz A, Engels S, Przyborski JM (2017) Proteomic analysis of exported chaperone/co-chaperone complexes of P. falciparum reveals an array of complex protein-protein interactions. Sci Rep 7:42188